

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 29, 2003, 20:49:12 ; Search time 330.604 Seconds
(without alignments)
8229.920 Million cell updates/sec

Title: US-09-988-971-1_COPY_517_684

Perfect score: 168

Sequence: 1 gccacacgcgctgcctcgg9.....gccacacgcgccaagtc 168

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: em_eelba:*
2: em_eelbm:*
3: em_eelbm:*
4: em_eelbm:*
5: em_eelbm:*
6: em_eelbm:*
7: em_eelbm:*
8: em_eelbm:*
9: gb_eel1:*
10: gb_eel2:*
11: gb_eel3:*
12: gb_eel4:*
13: gb_eel5:*
14: gb_eel6:*
15: em_eelbm:*
16: em_eelbm:*
17: gb_eel7:*
18: em_gse_hum:*
19: em_gse_inv:*
20: em_gse_pln:*
21: em_gse_vtc:*
22: em_gse_fun:*
23: em_gse_man:*
24: em_gse_mus:*
25: em_gse_other:*
26: em_gse_pro:*
27: em_gse_rtd:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match Length	ID	Description
1	168	100.0	778	BG178487 602328305
2	168	100.0	877	AL541041 AL541041
3	168	100.0	986	BQ054265 AGENCOURT
4	168	100.0	1002	BQ052308 AGENCOURT
5	168	100.0	1020	BQ054281 AGENCOURT
6	168	100.0	1069	BQ052468 AGENCOURT

7	161.6	96.2	878	14	BQ053486	BQ053486 AGENCOURT
8	124.8	74.3	627	10	BB619854	BB619854
9	124.8	74.3	640	14	BQ053006	BQ053006 H4019502-
10	124.8	74.3	660	10	BB635615	BB635615
11	123.8	73.7	670	9	A1510095	A1510095 m43c04.y
12	114.4	68.1	377	9	AA959151	AA959151 v251906.x
13	111	66.1	926	11	AK020837	AK020837 Mus muscu
14	90.8	54.0	506	17	B45150	B45150 HS-1060-B1-
15	83	49.4	541	13	BI988999	BI988999 480839 MA
16	67.2	40.0	389	10	BE015229	BE015229 127457 MA
17	52.4	31.2	525	13	BI348152	BI348152 da125c05.
18	50.4	30.0	507	9	AA051252	AA051252 m43c04.x
19	43.6	26.0	527	10	AA469655	AA469655 h432d07.x
20	43.6	26.0	546	10	BE397325	BE397325 601288664
21	43.6	26.0	559	10	AA968483	AA968483 E97380559
22	43.6	26.0	582	10	AA468724	AA468724 h437a06.x
23	43.6	26.0	653	12	BC389203	BC389203 602413834
24	43.6	26.0	689	10	AM303854	AM303854 xv21h02.x
25	43.6	26.0	698	12	BE890610	BE890610 601431488
26	43.6	26.0	716	10	BE676457	BE676457 7130c12.x
27	43.6	26.0	720	10	AM188088	AM188088 xj92c01.x
28	43.6	26.0	751	12	BF448700	BF448700 7n33c09.x
29	43.6	26.0	847	14	AM134909	AM134909 AU134909
30	43.6	26.0	847	14	BO682902	BO682902 AGENCOURT
31	43.6	26.0	953	14	BQ642677	BQ642677 AGENCOURT
32	43.6	26.0	963	14	BQ066964	BQ066964 AGENCOURT
33	43.6	26.0	1003	14	BQ062336	BQ062336 AGENCOURT
34	43.6	26.0	1006	14	BQ058222	BQ058222 AGENCOURT
35	43.6	26.0	1047	14	BQ066753	BQ066753 AGENCOURT
36	43.6	26.0	1051	14	BQ063502	BQ063502 AGENCOURT
37	43.6	26.0	1070	14	BM909567	BM909567 AGENCOURT
38	42.6	25.4	708	10	BE675166	BE675166 7f03f06.x
39	42	25.0	619	9	A1572989	A1572989 t64b01.x
40	41.6	24.8	340	10	AM481551	AM481551 36629 MAR
41	40.6	24.2	730	12	EG756039	EG756039 602716593
42	40.4	24.0	424	14	BQ305148	BQ305148 MR0-BT250
43	40	23.8	679	13	BI151771	BI151771 602915935
44	40	23.8	767	14	BM944156	BM944156 UI-M-BH0P
45	40	23.8	983	10	BB609447	BB609447

ALIGNMENTS

RESULT 1
BG178487 778 bp mRNA linear EST 06-FEB-2001
LOCUS 602328305F1 NIH_MGC_91 Homo sapiens CDNA clone IMAGE:4429896 5',
DEFINITION mRNA sequence.
ACCESSION BG178487
VERSION BG178487.1 GI:12685190
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 778)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-romail.nih.gov
Tissue Procurement: DCTD/DTP
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLNL0182 row: 1 column: 01
High quality sequence stop: 657.
location/Qualifiers
1. 778

FEATURES

source

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4429896"
 /clone_1ib="NIH MGC 91"
 /issue_type="adenocarcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: prostate; Vector: pCMVSPORT6, Site_1: NotI; Site_2: SalI; Cloned unidirectionally; oligo-dt primed. Average insert size 1.4 Kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH MGC Library."
 BASE COUNT 179 a 224 c 230 g 145 t
 ORIGIN

Query Match 100.0%; Score 168; DB 12; Length 778;
 Best Local Similarity 100.0%; Pred. No. 6.3e-35;
 Matches 168; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACAGCCGTGGCCCTGGGAGATTTCCTCCGAGAGTGGCCGCGAGCTGTCTGAGAGA 60
 DB 463 GCCACAGCCGTGGCCCTGGGAGATTTCCTCCGAGAGTGGCCGCGAGCTGTCTGAGAGA 522
 QY 61 CTCGGGAGCCATTGACCATGCTCTGAGAGATGAGACTGTGAGCGTGTCTGAA 120
 DB 523 CTCGGGAGCCATTGACCATGCTCTGAGAGATGAGACTGTGAGCGTGTCTGAA 582
 QY 121 GTCTCAGGAGAGATTAACATCCCGAGCTCCAGTGGCCAAAGTC 168
 DB 583 GTCTCAGGAGAGATTAACATCCCGAGCTCCAGTGGCCAAAGTC 630

RESULT 2
 AL541041 877 bp mRNA linear EST 16-FEB-2001
 LOCUS AL541041 LRT_FL002_PL1 Homo sapiens CDNA clone CS0DE005YK23 5 prime
 DEFINITION AL541041 LRT_FL002_PL1 Homo sapiens CDNA clone CS0DE005YK23 5 prime
 ACCESSION AL541041
 VERSION AL541041.1 GI:12871733
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 877)
 AUTHORS Li, W.-B., Gruber, C., Jessee, J., and Polayes, D.
 TITLE Full-length CDNA libraries and normalization
 JOURNAL Unpublished (2001)
 COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 Evry cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
 source
 1..877
 Location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="CS0DE005YK23"
 /clone_1ib="LRT_FL002_PL1"
 /lab_host="DH10B"
 /note="Organ: placenta; Vector: pCMVSPORT 6; 1st strand
 cDNA was primed with a NotI-oligo(dT) primer. Five prime
 end enriched, double-stranded cDNA was digested with Not I
 and cloned into the Not I and Eco RV sites of the
 pCMVSPORT 6 vector. Library was constructed by Life
 Technologies. Contact: Feng Liang Life Technologies, a
 division of Invitrogen 9800 Medical Center Drive Rockville
 , Maryland 20850, USA Fax : (1) 301 610 8371 Email :
 Eliang@lifetech.com URL :
 http://fulllength.invitrogen.com"

BASE COUNT 192 a 262 c 251 g 170 t 2 others
 ORIGIN

Query Match 100.0%; Score 168; DB 9; Length 877;
 Best Local Similarity 100.0%; Pred. No. 6.6e-35;

Matches 168; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GCCACAGCCGTGGCCCTGGGAGATTTCCTCCGAGAGTGGCCGCGAGCTGTCTGAGAGA 60
 DB 469 GCCACAGCCGTGGCCCTGGGAGATTTCCTCCGAGAGTGGCCGCGAGCTGTCTGAGAGA 548
 QY 61 CTCGGGAGCCATTGACCATGCTCTGAGAGATGAGACTGTGAGCGTGTCTGAA 120
 DB 549 CTCGGGAGCCATTGACCATGCTCTGAGAGATGAGACTGTGAGCGTGTCTGAA 608
 QY 121 GTCTCAGGAGAGATTAACATCCCGAGCTCCAGTGGCCAAAGTC 168
 DB 609 GTCTCAGGAGAGATTAACATCCCGAGCTCCAGTGGCCAAAGTC 656

RESULT 3
 BQ054265 966 bp mRNA linear EST 29-MAR-2002
 LOCUS BQ054265
 DEFINITION AGENCOURT 6830248 NIH_MGC_106 Homo sapiens CDNA clone IMAGE:5936339
 5', mRNA sequence.
 ACCESSION BQ054265
 VERSION BQ054265.1 GI:19813605
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 966)
 AUTHORS NIH-MGC http://mgs.nci.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cga@bds-remail.nih.gov
 Tissue Procurement: Dr. Daniel McVicar, DBS/NCI
 CDNA Library Preparation: Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.lnl.gov
 Plate: LNCM2125 row: 1 column: 12
 High quality sequence scop: 515.
 Location/Qualifiers
 1..966
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5936339"
 /clone_1ib="NIH MGC 106"
 /issue_type="natural killer cells, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: blood; Vector: pOTB7; Site_1: XhoI; Site_2:
 EcoRI; cDNA made by oligo-dT priming. Directionally cloned
 into EcoRI/XhoI sites using the following 5' adaptor:
 GGACAGG(G). Library constructed by Ling Hong in the
 laboratory of Gerald M. Rubin (University of California,
 Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
 Superscript II RT (Life Technologies). Note: this is a
 NIH MGC Library."

FEATURES
 source
 1..966
 Location/Qualifiers

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5936339"
 /clone_1ib="NIH MGC 106"
 /issue_type="natural killer cells, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: blood; Vector: pOTB7; Site_1: XhoI; Site_2:
 EcoRI; cDNA made by oligo-dT priming. Directionally cloned
 into EcoRI/XhoI sites using the following 5' adaptor:
 GGACAGG(G). Library constructed by Ling Hong in the
 laboratory of Gerald M. Rubin (University of California,
 Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
 Superscript II RT (Life Technologies). Note: this is a
 NIH MGC Library."
 BASE COUNT 211 a 291 c 276 g 207 t 1 others
 ORIGIN

Query Match 100.0%; Score 168; DB 14; Length 986;
 Best Local Similarity 100.0%; Pred. No. 6.9e-39;
 Matches 168; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACAGCCGTGGCCCTGGGAGATTTCCTCCGAGAGTGGCCGCGAGCTGTCTGAGAGA 60
 DB 361 GCCACAGCCGTGGCCCTGGGAGATTTCCTCCGAGAGTGGCCGCGAGCTGTCTGAGAGA 440
 QY 61 CTCGGGAGCCATTGACCATGCTCTGAGAGATGAGACTGTGAGCGTGTCTGAA 120
 DB 441 CTCGGGAGCCATTGACCATGCTCTGAGAGATGAGACTGTGAGCGTGTCTGAA 500

QY 121 GTCTCAGGAGAGATTAACATCCACCTCCACGTGGCCAAATC 168
 DB 501 GTCTCAGGAGAGATTAACATCCACCTCCACGTGGCCAAATC 548

RESULT 4
 LOCUS BQ052308
 DEFINITION AGENCOURT_6868571 NIH_MGC_106 Homo sapiens cDNA clone IMAGE:5933542
 5', mRNA sequence.
 ACCESSION BQ052308
 VERSION BQ052308
 KEYWORDS BQ052308.1 GI:19811648
 SOURCE EST.
 ORGANISM human.
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 1002)
 NIH-MGC http://mgs.nci.nih.gov/
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cga@bbs-remail.nih.gov
 Tissue Procurement: Dr. Daniel McVicar, DBS/NCI
 cDNA Library Preparation: Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: LCM218 row: d column: 23
 High quality sequence stop: 670.

FEATURES

source
 Location/Qualifiers
 1..1002
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5933542"
 /clone_1b="NIH MGC 106"
 /tissue_type="natural killer cells, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: blood; Vector: pOTB7; Site 1: XhoI; Site 2:
 EcoRI; cDNA made by oligo-dT priming. Directionally cloned
 into EcoRI/XhoI sites using the following 5' adaptor:
 GGACAGAG(G). Library constructed by Ling Hong in the
 laboratory of Gerald M. Rubin (University of California,
 Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
 Superscript II RT (Life Technologies). Note: this is a
 NIH MGC library."

BASE COUNT 221 a 296 c 288 g 197 t
 ORIGIN

Query Match 100.0%; Score 168; DB 14; Length 1002;
 Best Local Similarity 100.0%; Pred. No. 7e-39;
 Matches 168; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACAGCCGCGCCCTGGGAGATTCCCGGAGGTGCGCGGCGAGTGCCTGAGA 60
 DB 198 GCCACAGCCGCGCCCTGGGAGATTCCCGGAGGTGCGCGGCGAGTGCCTGAGA 257
 QY 61 CTCGGGAGACCATTTGACCATCTCTTGAGAGATGAGACTGTGAGACGGTGTCTGAA 120
 DB 258 CTCGGGAGACCATTTGACCATCTCTTGAGAGATGAGACTGTGAGACGGTGTCTGAA 317
 QY 121 GTCTCAGGAGAGATTAACATCCACCTCCACGTGGCCAAATC 168
 DB 318 GTCTCAGGAGAGATTAACATCCACCTCCACGTGGCCAAATC 365

RESULT 5
 LOCUS BQ054281
 DEFINITION AGENCOURT_6830234 NIH_MGC_106 Homo sapiens cDNA clone IMAGE:5936362
 5', mRNA sequence.

ACCESSION BQ054281
 VERSION BQ054281.1 GI:19813621
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 1020)
 NIH-MGC http://mgs.nci.nih.gov/
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cga@bbs-remail.nih.gov
 Tissue Procurement: Dr. Daniel McVicar, DBS/NCI
 cDNA Library Preparation: Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: LCM2125 row: j column: 11
 High quality sequence stop: 556.

FEATURES
 source
 Location/Qualifiers
 1..1020
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5936362"
 /clone_1b="NIH MGC 106"
 /tissue_type="natural killer cells, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: blood; Vector: pOTB7; Site 1: XhoI; Site 2:
 EcoRI; cDNA made by oligo-dT priming. Directionally cloned
 into EcoRI/XhoI sites using the following 5' adaptor:
 GGACAGAG(G). Library constructed by Ling Hong in the
 laboratory of Gerald M. Rubin (University of California,
 Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
 Superscript II RT (Life Technologies). Note: this is a
 NIH MGC library."

BASE COUNT 219 a 311 c 283 g 204 t 3 others
 ORIGIN

Query Match 100.0%; Score 168; DB 14; Length 1020;
 Best Local Similarity 100.0%; Pred. No. 7e-39;
 Matches 168; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACAGCCGCGCCCTGGGAGATTCCCGGAGGTGCGCGGCGAGTGCCTGAGA 60
 DB 314 GCCACAGCCGCGCCCTGGGAGATTCCCGGAGGTGCGCGGCGAGTGCCTGAGA 373
 QY 61 CTCGGGAGACCATTTGACCATCTCTTGAGAGATGAGACTGTGAGACGGTGTCTGAA 120
 DB 374 CTCGGGAGACCATTTGACCATCTCTTGAGAGATGAGACTGTGAGACGGTGTCTGAA 433
 QY 121 GTCTCAGGAGAGATTAACATCCACCTCCACGTGGCCAAATC 168
 DB 434 GTCTCAGGAGAGATTAACATCCACCTCCACGTGGCCAAATC 481

RESULT 6
 LOCUS BQ052468
 DEFINITION AGENCOURT_6868422 NIH_MGC_106 Homo sapiens cDNA clone IMAGE:5933772
 5', mRNA sequence.
 ACCESSION BQ052468
 VERSION BQ052468
 KEYWORDS BQ052468.1 GI:19811808
 SOURCE EST.
 ORGANISM human.
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 1069)
 NIH-MGC http://mgs.nci.nih.gov/
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)

JOURNAL

Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: c9apbs-remail.nih.gov

Tissue Procurement: Dr. Daniel McVicar, DMS/NCI
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLN at:
http://image.llnl.gov

Plate: LNCM2118 row: n column: 13
High quality sequence stop: 681.

FEATURES

source

Location/Qualifiers
1. .878
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5935253"
/clone_1lb="NIH-MGC-106"
/tissue_type="natural killer cells, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: blood; Vector: pOT7; Site: 1: XhoI; Site 2:
EcoRI; CDNA made by oligo-dt priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGGAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH-MGC Library."

FEATURES

source

Location/Qualifiers
1. .1069
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5933772"
/clone_1lb="NIH-MGC-106"
/tissue_type="natural killer cells, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: blood; Vector: pOT7; Site: 1: XhoI; Site 2:
EcoRI; CDNA made by oligo-dt priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGGAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH-MGC Library."

BASE COUNT

230 a 328 c 300 g 205 t 6 others

ORIGIN

Query Match 100.0%; Score 168; DB 14; Length 1069;
Best Local Similarity 100.0%; Pred. No. 7.2e-39;
Matches 168; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACAGCCGTCGCTGCGGAGCTTCCCGGAGGTGCGCCGCGAGCTGTCTGAGA 60
DB 191 GCCACAGCCGTCGCTGCGGAGCTTCCCGGAGGTGCGCCGCGAGCTGTCTGAGA 250
QY 61 CTCGGGGAGCATTGACATGCTCTGAGAGTGAAGTGGTGGAGCGGTCTGTGAA 120
DB 251 CTCGGGGAGCATTGACATGCTCTGAGAGTGAAGTGGTGGAGCGGTCTGTGAA 310
QY 121 GTCTCAGGCGAGAGATTAACATCCCGAGCTCCACGTGGCCAAAGTC 168
DB 311 GTCTCAGGCGAGAGATTAACATCCCGAGCTCCACGTGGCCAAAGTC 358

RESULT 7

LOCUS

B0053486 878 bp mRNA linear EST 29-MAR-2002

DEFINITION

AGENCOURT_6822017 NIH-MGC_106 Homo sapiens cDNA clone IMAGE:5935253

ACCESSION

B0053486 5', mRNA sequence.

VERSION

B0053486.1 GI:19812826

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens

REFERENCE

1 (bases 1 to 878)

AUTHORS

NIH-MGC http://imgc.nci.nih.gov/.

TITLE

National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL

Unpublished (1999)

COMMENT

Contact: Robert Strausberg, Ph.D.

REFERENCE

1 (bases 1 to 878)

AUTHORS

NIH-MGC http://imgc.nci.nih.gov/.

TITLE

National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL

Unpublished (1999)

COMMENT

Contact: Robert Strausberg, Ph.D.

REFERENCE

1 (bases 1 to 878)

AUTHORS

NIH-MGC http://imgc.nci.nih.gov/.

TITLE

National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL

Unpublished (1999)

COMMENT

Contact: Robert Strausberg, Ph.D.

Plate: LNCM2122 row: 1 column: 06
High quality sequence stop: 394.

FEATURES

source

Location/Qualifiers
1. .878
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5935253"
/clone_1lb="NIH-MGC-106"
/tissue_type="natural killer cells, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: blood; Vector: pOT7; Site: 1: XhoI; Site 2:
EcoRI; CDNA made by oligo-dt priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGGAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH-MGC Library."

BASE COUNT

201 a 253 c 233 g 190 t 1 others

ORIGIN

Query Match 96.2%; Score 161.6; DB 14; Length 878;
Best Local Similarity 97.6%; Pred. No. 5e-37;
Matches 164; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GCCACAGCCGTCGCTGCGGAGCTTCCCGGAGGTGCGCCGCGAGCTGTCTGAGA 60
DB 375 GCCACAGCCGTCGCTGCGGAGCTTCCCGGAGGTGCGCCGCGAGCTGTCTGAGA 434
QY 61 CTCGGGGAGCATTGACATGCTCTGAGAGTGAAGTGGTGGAGCGGTCTGTGAA 120
DB 435 CTCGGGGAGCATTGACATGCTCTGAGAGTGAAGTGGTGGAGCGGTCTGTGAA 494
QY 121 GTCTCAGGCGAGAGATTAACATCCCGAGCTCCACGTGGCCAAAGTC 168
DB 495 ATCTCAGGCGAGAGATTAACATCCCGAGCTCCACGTGGCCAAAGTC 542

RESULT 8

LOCUS

B619854 627 bp mRNA linear EST 31-AUG-2001

DEFINITION

BB619854 RIKEN full-length enriched, adult male thymus Mus musculus

ACCESSION

BB619854 CDNA clone 5930437K10 5', mRNA sequence.

VERSION

BB619854.1 GI:15396929

KEYWORDS

EST.

SOURCE

house mouse.

ORGANISM

Mus musculus

REFERENCE

1 (bases 1 to 627)

AUTHORS

Arakawa, T., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A.,

TITLE

Arakawa, T., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A.,

JOURNAL

Arakawa, T., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A.,

COMMENT

Arakawa, T., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A.,

REFERENCE

Arakawa, T., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A.,

AUTHORS

Arakawa, T., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A.,

TITLE

Arakawa, T., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A.,

JOURNAL

Arakawa, T., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A.,

COMMENT

Arakawa, T., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A.,

REFERENCE

Arakawa, T., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A.,

AUTHORS

Arakawa, T., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A.,

TITLE

Arakawa, T., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A.,

JOURNAL

Arakawa, T., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A.,

COMMENT

Arakawa, T., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A.,

Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Knehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G., Quackenbush, J., Schriml, L.M., Stambli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barish, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H., Toyooka, K., Wang, K.H., Weitz, C., Whitaker, C., Wilming, L., Wyshaw-Borje, A., Yoshida, K., Haegawa, Y., Kawaji, H., Kohenki, S., and Hayashizaki, Y.

Functional annotation of a full-length mouse cDNA collection
Nature 409 (6821), 685-690 (2001)

21085660
11217851

5 (bases 1 to 926)
Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A., Arawaka, T., Baldarelli, R., Bono, H., Brownstein, M., Bult, C., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hangaki, T., Hara, A., Hayatsu, N., Hill, D., Hiramoto, K., Hirooka, T., Hori, F., Hume, D., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kabukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Quackenbush, J., Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, H., Schriml, L., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Tejima, Y., Toya, T., Yamamura, T., Yamanaka, I., Yanushkin, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.

Direct Submission
Submitted (18-AUG-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Shihito-cho, Tsukuba, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-resgsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

COMMENT
Please visit our web site (http://genome.gsc.riken.go.jp/) for further details.
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5' GAGAGAGAGATTCGAGATTCGAGAGCTTTTCTTTTCTTTTNN 3'], cDNA was prepared by using triazole thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through two rounds of normalization to Rot = 20.0 and subtraction to Rot = 458.8. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGATTCGAGATTCGAGATTCGAGAGCTTTTCTTTTCTTTTNN 3']. cDNA was cleaved with BamHI and XhoI. Vector: a modified plasmid, SC(+) after bulk excision from Lambda phage. Cloning sites: 5' end: SalI, 3' end: BamHI. Host: DH10B. Retina RNA was provided by Stefano Gustincich (Department of Neurobiology, Harvard Medical School, 220 Longwood Ave., Boston, MA 02115, USA) whose assistance is gratefully acknowledged.

FEATURES

Source

1. 926
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="PANTOM_DB:A930009E21"
/db_xref="MGD:MGI:1911678"
/db_xref="taxon:10090"
/clone="A930009E21"

CDS

/cissue_type="retina"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
128..667
/note="data source: SPT, source key: 060898, evidence: ISS putative
similar to SRC-LIKE ADAPTER PROTEIN"
/protein_id="AB32223.1"
/db_xref="GI:12861543"
/translation="MPSVYAKAHGWYELSEKAEILLIPNPGALIREST
RRCVSLVRLSRPASWDRIRHRYIQRDNMVLVSRITFPLHVLVHSELDADI
CPRLRPVLOGLGPGKTPPVVPTSSINMKLDRSLFLLEAPASGEASLSBG
LRESYSYSLAEDPLDPA"
907..912
/note="putative"

polyA_signal
polyA_site
926
/note="putative"

BASE COUNT 209 a 269 c 263 g 185 t
ORIGIN

Query Match 66.1%; Score 111; DB 11; Length 926;
Best Local Similarity 83.4%; Pred. No. 3.6e-22;
Matches 126; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

QY 18 GGGCAATTTCCCGGAGAGTGGCCCGGAGCTGTGAGATTCGGGGAGCCATTGAC 77

DB 4 GGGCAATTTCCCGGAGAGTGGCCCGGAGCTGTGAGATTCGGGGAGCCATTGAC 63

QY 78 CATGCTCTGAGAGTGGAGAGTGTGACGAGTGTCTGAGATTCGAGAGAGTA 137

DB 64 CATCATCTCTGAGAGTGGAGAGTGTGACGAGTGTCTGAGATTCGAGAGAGTA 123

QY 138 TAACATCCCGAGCTCCACGCTGCGCAAAAGTC 168

DB 124 CCAATGCCCAAGTGTATGAGCTAAAGTC 154

RESULT 14
B45150
LOCUS

DEFINITION HS-1060-B1-H07-WF.abi 506 bp DNA linear GSS 21-OCT-1997

ACCESSION B45150
VERSION B45150.1 GI:2549984

KEYWORDS GSS.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

AUTHORS 1 (bases 1 to 506)
Mahairas, G.G., Zackrone, K.D., Smith, T., Tipton, S., Schindl, S.,
Tyalcoff, R., Abajian, C., Blanchard, A., West, A. and Hood, L.B.

TITLE Construction of a Characterized Clone Resource for Genomic
Sequencing: Generation and Preliminary Analysis of 20,000 Sequence
Tagged Connectors

JOURNAL Unpublished (1997)
COMMENT Contract: Mahairas GG, Zackrone KD, Hood L
University of Washington
Seattle, WA 98195, USA

TELE (206) 616-8744
FAX: (206) 685-7301
Email: kzackron@u.washington.edu

Sequence Tagged Connector
Plate: CT 782 row: P column: 13

Class: BAC ends
High quality sequence strop: 506.

Location/Qualifiers

1. 506
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="Plate=CT 782 Col=13 Row=P"


```

/clone_lib="CIT Human Genomic Sperm Library C"
/bex="NM"
/notes="Organ: sperm; Vector: pBel0BAC11; BAC Clones in
E-Coli DH10B"
BASE COUNT      105 a      128 c      161 g      112 t
ORIGIN

```

```

Query Match      54.0%; Score 90.8; DB 17; Length 506;
Best Local Similarity 93.1%; Pred. No. 2.4e-16;
Matches 95; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```

```

QY 1 GCCACAGCCGTCGCTGGGAGTTTCCCGGAGGTGCCCCGCGAGCTGTGCTGAGA 60
DB 223 GCCACAGCCGTCGCTGGGAGTTTCCCGGAGGTGCCCCGCGAGCTGTGCTGAGA 282
QY 61 CTGGGGAGCCATTGACATTCGTCTGAGATGAGACTGTGAGAC 107
DB 283 CTGGGGAGCCATTGACATTCGTCTGAGATGAGACTGTGAGAC 324

```

```

RESULT 15
BI989899 541 bp mRNA linear EST 16-OCT-2001
LOCUS 480839 MARC 2BOV Bos taurus cDNA 5', mRNA sequence.
ACCESSION BI989899
VERSION BI989899.1 GI:16187055
KEYWORDS EST.
SOURCE
  Bos taurus
  cow.
  ORGANISM
    Eukaryota; Metazoa; Chordata; Cranialia; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;
    Bovidae; Bovinae; Bos.

```

```

REFERENCE
AUTHORS Smith,T.P.L., Grose,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,
  Cabasa,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C., Bennett,
  G.L., Heaton,W.P., Laegreid,W.W., Rohrer,G.A., Chitko-McKown,C.G.,
  Perlea,G., Holt,I., Karamycheva,S., Liang,F., Quackenbush,J. and
  Keele,J.W.
  Sequence evaluation of four pooled-tissue normalized bovine cDNA
  libraries and construction of a gene index for cattle
  Genome Res. 11 (4), 626-630 (2001)
  21180013
  Contact: Smith TPL
  USDA, ARS, US Meat Animal Research Center
  PO Box 166, Clay Center, NE 68933-0166, USA
  Tel: 402 762 4366
  Fax: 402 762 4390
  Email: smith@email.marc.usda.gov
  Single pass sequencing. Bases called and alt trimmed with phred
  v0.980904.e. Vector identified by cross_match with the -minscore 18
  and -minmatch 12 options.
  PCR Primers
  FORWARD: AGGAAACAGCTAGACAT
  BACKWARD: GTTTCACAGTCAGACG
  Plate: 105 row: P column: 20
  Seq primer: ATTAGTGACACTATAG.

```

```

FEATURES
  source
    1..541
      /organism="Bos taurus"
      /db_xref="taxon:9913"
      /clone_lib="MARC 2BOV"
      /issue_type="pooled"
      /lab_host="DH10B"
      /note="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;
      library made from pooled tissue from testis, thymus,
      semitendinosus muscle, longissimus muscle, pancreas,
      adrenal, and endometrium."

```

```

BASE COUNT      120 a      160 c      157 g      104 t
ORIGIN

```

```

Query Match      49.4%; Score 83; DB 13; Length 541;
Best Local Similarity 86.0%; Pred. No. 4.8e-14;
Matches 92; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

```

```

QY 1 GCCACAGCCGTCGCTGGGAGTTTCCCGGAGGTGCCCCGCGAGCTGTGCTGAGA 60
DB 435 GCCAAGCCGTCGCTGGGAGTTTCCCGGAGGTGCCCCGCGAGCTGTGCTGAGA 494
QY 61 CTGGGGAGCCATTGACATTCGTCTGAGATGAGACTGTGAGAC 107
DB 495 CTGGGGAGCCATTGACATTCGTCTGAGATGAGACTGTGAGAC 541

```

```

Search Completed: March 30, 2003, 02:49:21
Job time : 334.604 secs

```

